HEPEX, the Hydrological Ensemble Prediction Experiment

John C. Schaake

Consultant, NOAA National Weather Service, Office of Hydrologic Development

Silver Spring, Maryland

Thomas M. Hamill

NOAA Earth System Research Laboratory, Physical Sciences Division

Boulder, Colorado

Roberto Buizza

European Centre for Medium Range Weather Forecasts

Reading, England

Martyn Clark

National Institute for Water and Atmospheric Research (NIWA)

Christchurch, New Zealand

Revised, 4 January 2007

Submitted to *Bulletin of the American Meteorological Society*

Corresponding author address:

Dr. John C. Schaake
1A3 Spa Creek Landing
Annapolis, MD 21403
e-mail: john.schaake@noaa.gov
phone: (410) 320-4838

ABSTRACT

The Hydrological Ensemble Prediction Experiment (HEPEX) is an international project to advance technologies for hydrological forecasting. Its goal is "*to bring the international hydrological and meteorological communities together to demonstrate how to produce and utilize reliable hydrological ensemble forecasts to make decisions for the benefit of public health and safety, the economy, and the environment*." HEPEX is an open group comprised primarily of researchers, forecasters, water managers, and users. HEPEX welcomes new members.

In the first workshop, held in the spring of 2004, HEPEX participants formulated scientific questions that, once addressed, should help produce valuable hydrological ensemble prediction to serve users' needs. During the second HEPEX workshop, held in the summer of 2005, a series of coordinated test-bed demonstration projects was set up as a method for answering these questions. The test beds are collections of data and models for specific hydrological basins or sub-basins, where relevant meteorological and hydrological data has been archived. The test beds will facilitate the inter-comparison of various hydrological prediction methods and linkages to users. The next steps for HEPEX are to complete the work planned for each test bed and to use the results to engineer more valuable automated hydrological prediction systems.

1. **A rationale for hydrological ensemble prediction**

Imagine yourself as the manager of a reservoir in the western USA. Finally, after many years of drought and low water levels, the mountains above you have received ample snowfall this winter. It's late spring now, and the extended-range forecast suggests a strong surge of moisture. A single forecast based on a (possibly high-resolution) model prediction indicates heavy rain on the snow pack, causing very rapid melting, perhaps producing more flow than your reservoir can store. If you release water from the reservoir now in anticipation of extreme runoff and the precipitation is less than predicted, that water is lost to your customers; should the drought return, inadequate reservoir storage may eventually require water rationing. But if you don't release, there's a chance that the sudden surge of water could top the reservoir and cause potentially catastrophic flooding downstream.

This is an example of one of many complex decisions faced by water managers. Ideally, as manager, you would be supplied with a perfect weather forecast, you'd have precise measurements of the snow pack and soil moisture, and you'd utilize highly engineered hydrological models that would near perfectly predict the amount and timing of the streamflow. The one resulting hydrological prediction would provide enough information to make the correct decision. In reality, there are tremendous uncertainties. The weather forecasts supplied to you are imperfect and lacking in critical detail; will the precipitation fall primarily in the form of rain on snow (bad, as it may cause flash flooding) versus snow on snow (good, as it would generate a gradual, delayed runoff)? At what elevation will the rain change to snow? And what about the existing snow pack? There may be only a handful of actual snow depth measurements. Finally, the land-surface and hydrological routing models you have available are commonly simplified descriptions of the hydrological processes; for example, they may treat each sub-basin as a homogeneous element covered by the same average snow cover and soil moisture. Given the myriad of uncertainties, a natural tool for making the decisions would be a probabilistic forecast, possibly based on an ensemble hydrological prediction system, akin to the now ubiquitous ensemble weather prediction systems (Buizza et al. 2005). Ideally, this system would produce multiple realizations of possible future streamflows that were "sharp" (much more specific than, say, drawing from a climatology of streamflows in past years) and yet reliable (e.g., over many situations, when there was a 20 percent forecast of a runoff exceeding y m^3s^{-1}, the runoff actually exceeded y m^3s^{-1} 20 percent of the time). Were such a product available, the eventual cost of reduced storage from a dam release could be weighed against the likelihood of flooding impacts without the release.

An automated, skillful, reliable ensemble streamflow forecast product is conceptually appealing. Figure 1 provides a schematic of one possible system that explicitly accounts for the major sources of uncertainty in the forecasting process. An ensemble of atmospheric forecasts is first run through a meteorological pre-processor, producing meteorological forcings for the hydrological model that have been downscaled, corrected for bias, converted to produce the specific variables of interest, and adjusted to have realistic spatial and temporal correlations of errors. Meanwhile, all the available measurements of soil moisture, snow depth and water equivalent, and even perhaps satellite and radar data have been utilized in an ensemble hydrological data assimilation system. This system produces an ensemble of plausible analyses of the state of the land surface, snow pack, and initial stream flow, all with realistic spatial correlation of errors. Land-surface and hydrological routing models, or perhaps an ensemble of models are now run, coded so that any possible deficiencies in the models will realistically increase the spread of possible outcomes (e.g., by using multiple feasible parameter sets for each hydrological model). A hydrological product generator is run to correct for remaining systematic errors and translate the output into the formats and variables of interest for particular users. This output is evaluated by users and objectively verified. Calibrated probabilistic forecast systems such as this were recently recommended for widespread adoption by the National Research Council (2006).

End-to-end systems like this are just beginning to be assembled, and there are many basic and applied science questions that must be answered in order to build useful systems. HEPEX, the *Hydrological Ensemble Prediction Experiment*, is a project specifically designed by hydrologists, meteorologists, and users to answer these questions and promote the rapid development of such systems. HEPEX was launched in March 2004 at a meeting at the European Centre for Medium Range Weather Forecasts (ECMWF). Since that workshop, HEPEX has organized sessions on ensemble prediction at various international conferences, and a second HEPEX workshop was recently held in July 2005 at the National Center for Atmospheric Research (NCAR) in Boulder, Colorado). The official HEPEX web page is http://hydis8.eng.uci.edu/hepex/ .

The intent of this publication is to explain HEPEX to the community and to encourage others with relevant experience to participate. HEPEX is an open, participatory project. It is not directly funded by any agency, but rather shaped in a bottom-up process by scientists and users who strongly believe that improved forecast techniques can be built most effectively through interdisciplinary collaboration. In the rest of the article, we address the following questions: (1) what are the motivating science questions? [section 2.] (2) What is our proposed method for answering these questions? [HEPEX test bed sidebar.] (3) What is the current status of the development of automated hydrological forecast systems? [another sidebar.] (4) How is HEPEX organized, and how does this it link with other hydrological research programs? [section 3.]

2. **HEPEX science questions.**

Working from the bottom of Fig. 1 upwards, let us consider some of the key scientific questions that HEPEX hopes to address associated with each major component of a probabilistic hydrological forecast system. The accompanying sidebar describes roughly where we are in 2007 in building systems to address these questions and produce high-quality, automated ensemble streamflow forecasts.

a. User issues

Who are the primary customers and potential customers of hydrological forecasts, and who else might use these forecasts if they were more skillful and reliable? How can we improve communication of scientific discoveries to the customers, and how can we tailor hydrological systems to meet their requirements? HEPEX has instituted a users' committee that will guide the research projects towards addressing key questions like these. While still more outreach is needed, this committee has some preliminary answers to these questions. Users of hydrological forecasts may include reservoir and city water-supply managers, emergency management professionals, and environmental scientists concerned about water quality or fisheries. Agriculture, insurance, navigation, and power-generation industries may want such products. For intensive users, HEPEX participants will seek to determine how the data can most effectively be presented in order to aid in their decision-making process (part of the envisioned hydrological product generator in Fig. 1). On the other hand, many customers may not be familiar with how they can optimize their decisions based on probabilistic information (e.g. Zhu et al. 2002), so HEPEX participants will also educate users and help them adapt their existing practices.

b. Hydrological forecast verification

After hydrological forecasts have been created, they should also be verified (Fig. 1, bottom right). Important scientific questions in hydrological system verification include, "what are the intrinsic limits of hydrological predictability for a given basin?" "How can we statistically evaluate the skill of our systems the most important extreme events, which are by nature rare?" "How do we quantify any value added by the human forecaster in this process?" "What verification information do users need?" and "how do we verify the many important aspects of the hydrological system, such as the spatial and temporal correlations of the input forecasts?"

Some obvious steps will foster improved hydrological verification. Routine measurements of precipitation, soil moisture, snow cover, streamflow, and other related hydrological variables should be processed and stored in common formats so that all components of the system can be verified. Additional data may include customer-specific measurements such as the amount of hydropower that was generated. Depending on the application, data may be needed at high spatial and temporal resolution (e.g., for flash floods) or accumulated over large areas and long periods (large river management). Since many rivers cross national borders, the international sharing of hydrological observations will aid the development of hydrological prediction systems.

Verification is difficult when the most interesting event is rare, be it a flash flood or the yearly peak runoff or the forecast during a rare extreme event. Testing a model across many basins is one way to increase the sample size of these rare events.
Also, to assess the statistical significance of changes to a hydrological forecast system, a long time series of streamflow hindcasts and observations will be helpful. Ideally, this then requires that prior weather forecasts from a consistent model should be available (Hamill et al. 2006).

c. Hydrological product generation

What are appropriate techniques for calibrating streamflow predictions? Despite the best efforts, an ensemble of streamflow simulations fed with an ensemble of realistic forcings and using state-of-the art hydrological models may still produce biased streamflow estimates. Hence, calibration of hydrological model output through a "hydrological product generator" is envisioned. Such a product generator would also reformat the data to be most convenient for users. For example, a reservoir operator may require quantitative streamflow forecasts from all ensemble members to feed into a decision support model to optimize releases of water from multiple reservoirs. Emergency managers may require graphical output that displays the probability that streamflow is above a prescribed threshold several hours into the future. And irrigators may require information on the number of days that streamflow will remain above a prescribed threshold and be available for their use. Each of these sources of information is available from the ensemble streamflow prediction system; hydrological product generation needs to be flexible enough to provide information for these myriad of uses.

d. Hydrological models.

In the HEPEX concept, an ensemble of hydrological models will be run, with different hydrological forecast members forced by different calibrated atmospheric forecasts and different but plausible initial soil, snow, and river conditions. The uncertainties in the hydrological models will themselves be accounted for explicitly. Broad science questions include "what are the sources of uncertainty in the hydrological forecast system?" "How do we formulate a hydrological system to account for all the effects of uncertainty?" "Can we quantify the relative contribution of each source of uncertainty upon the resulting hydrological forecast uncertainty?" And "what is the value of more complex, 'distributed' hydrological approaches relative to the more simplified, 'lumped' representations?"

Of course we know that the atmospheric forecasts are uncertain, but it is less clear whether these predominate over the uncertainties in the hydrological initial conditions and uncertainties in the hydrological model itself. The relative contributions of the uncertainties may vary from one situation to the next. For a flash flood, the primary uncertainty may be the weather forecast itself, but for springtime runoff in, say, the Colorado River, for example, the accuracy of streamflow forecasts may be controlled primarily by the accuracy of the mountain snow pack analysis.

e. Hydrological data assimilation.

Important questions include “what are the most appropriate methods for generating an ensemble of land-surface state estimates?”, “how can we incorporate new sources of data such as satellite radiances into systems primarily designed to assimilate ground-based observations?” and “to what extent are streamflow forecast skill increased through improved hydrological data assimilation techniques? Does the impact of improved techniques vary between catchments with fast and slow response times?” Also, “can we estimate of the sub-gridscale heterogeneity of the state given the observational data?”

The proposed ensemble hydrological forecast system requires an ensemble of plausible estimates of the current state of the land surface (e.g., soil moisture, ground water, and snow-water equivalent) and streamflow. This ensemble should have the property that its mean is a minimum-error estimate of the current state. The spread of the ensemble reflects the inherent analysis uncertainty given the scattered input data. Also, the covariances among state components should be properly modeled (e.g., nearby hilltops will have more positively correlated snow-depth estimates than will the hilltops and valleys). This complex, highly heterogeneous state must be inferred from widely scattered observations, and often the variables of interest such as soil moisture are not directly measured but must be modeled from a time series of temperature, wind, and estimated precipitation (e.g., Mitchell et al. 2004) or from proxy information such as near-surface humidity or satellite data (Seuffert et al. 2004). Data from atmospheric models may be coarse in time and space resolution, so that they may require a statistical downscaling.

f. Pre-processing atmospheric weather-climate forecasts.

HEPEX seeks to address several questions on how to optimally use meteorological ensemble predictions. "What are the requirements of weather-climate forecasts to support hydrological ensemble prediction, and do existing ensemble products meet them?" "What is the appropriate role of the human forecaster relative to machine-generated products?" "What is the value added from post-processed versus raw ensemble forecasts?" and "how do intermittent phenomena such as El Niño modulate the weather and climate forecasts?"

We know that the forecasts should be sharp yet reliable, and they should provide realistic, small-scale detail if the hydrological problem (e.g., flash-flood forecast) requires this. Unfortunately, we also know that ensemble predictions are often contaminated by model biases, and the observed weather too frequently lies outside the span of the ensemble (i.e., the ensemble spread should be larger). Also, the ensemble forecasts are conducted with reduced-resolution models, less capable to provide predictions with the required small-scale detail. Consequently, HEPEX envisions that pre-processing of ensemble weather and climate forecasts will be necessary to correct bias and spread errors and to downscale. This calibration may require the use of a large data set of forecasts from a fixed forecast model and data assimilation system.

SIDEBAR 1

The current state-of-the-art of hydrological ensemble streamflow prediction.

[insert into text near section 2]

How near is the hydrological community to having the end-to-end "Community Hydrologic Prediction System" (CHPS) like Fig. 1 in place?

Components of the atmospheric pre-processor that HEPEX envisions are now being developed. Hamill et al. (2006) have a freely accessible reforecast data set and real-time forecasts available, and have demonstrated how such a data set may be useful for calibrating precipitation and temperature forecasts. Similar data sets are being developed at ECMWF and Environment Canada. Multi-model reforecasts like those provided by the THe Observing system Research and Predictability Experiment (THORPEX) Interactive Grand Global Experiment (TIGGE) may also be valuable (TIGGE, 2005). To produce ensemble precipitation and temperature forcing for its Advanced Hydrologic Prediction Service (AHPS), the U.S. National Weather Service has developed an initial atmospheric forecast pre-processor (Schaake et al. 2006).

Hydrological data assimilation has improved significantly in the last decade with the operational implementation of land-state assimilation techniques at operational centers (e.g., Mitchell et al. 2004) and the development of advanced probabilistic techniques (Reichle et al. 2002, Seo et al. 2003, Slater and Clark 2006, Clark et al. 2006).

Regarding the development of hydrological forecast models, many groups have developed sophisticated land-surface models coupled to routing models (Bandaroga et al. 2004, Koren et al. 2004). Results from the Distributed Model Intercomparison Project (DMIP) address some of these questions in a special issue of the *Journal of Hydrology*

(Volume 298, October 2004), and Phase 2 of DMIP is under way to answer additional questions.

Very little research has been performed on hydrologic product generation methods. One initial effort that is being integrated into National Weather Service hydrological forecasting is recent work by Seo et al. (2006).

For hydrological verification, the quality of streamflow observations varies widely across the world. Some basins have a long time series of flow measurements, others do not. The HEPEX test beds (see sidebar) offer collections of streamflow and land-surface data for several basins, many of which have long records.

SIDEBAR 2:

Test beds, an important component for achieving the HEPEX goal.

[insert around sections 2-3]

Much of the effort in science is in the collection of data and the coding of models. A rational way to speed the development of hydrological forecast systems at low cost is through sharing these data and models. Shared results from the communal data can also be helpful for indicating whether a proposed method provides a general improvement upon existing procedures.

Accordingly, HEPEX plans to achieve hydrological ensemble forecast improvements through a series of test beds. Most of the HEPEX test beds have access to weather ensemble forecasts, associated observations, land-surface analyses, streamflow measurements, and ancillary information on uncertainty. Some test beds may include demonstration codes so that other researchers can compare results and may contribute to the future development of a CHPS. Eight HEPEX test beds have been selected so far (Table 1), but many more may be added. Six of them represent a variety of basins or sub-basins with different terrain, different climatologies, different hydrological issues, different data densities, and differences in the amount of regulation of stream flows in the basin. The remaining two test beds are collections of model codes to be inter-compared. Each test bed has one or more hosts, an investigator or institution responsible for gathering and maintaining the data and codes.

We envision that test beds constitute a natural framework to address many of the questions proposed in section 2. A detailed description of each test bed and the scientific questions to be answered are provided at the HEPEX web site, http://hydis8.eng.uci.edu/hepex/.

3. **HEPEX organization and affiliations**

HEPEX has two steering committees, a users committee and a scientific committee acting as the main coordinating bodies. The committee members represent a mixture of areas of expertise, geographical regions, and institutional capabilities. As needed, there will be sub-committees to address specific scientific issues such as data management, downscaling techniques, and these will interact with test-bed leaders. Scientists and users who are not yet part of HEPEX and who want to help forward its goals are encouraged to contact HEPEX organizers (John Schaake, John.Schaake@noaa.gov, and Roberto Buizza, Roberto.Buizza@ecmwf.int). Appointments to the committees will be revisited and revised at the 3rd HEPEX workshop planned for June, 2007.

HEPEX is a global project affiliated with several international organizations. The initial impetus for HEPEX grew out of a need to help the World Climate Research Program's (WCRP) Global Water and Energy Cycle Experiment (GEWEX) meet its water-resource applications objectives. The World Meteorological Organization's Hydrology and Water Resources Program (HWRP) is assisting HEPEX meet the needs of National Hydrological Services who will use HEPEX products. HEPEX expects the International Association of Hydrological Sciences (IAHS) Predictions for Ungaged Basins (PUB) initiative will contribute both new science and data sets, and will participate in some of the test bed projects. Ensemble atmospheric forecasts are expected to be available for HEPEX applications from a number of models participating in the World Weather Research Project's (WWRP) THORPEX Interactive Grand Global Ensemble Experiment (TIGGE). Finally, HEPEX is assisting the inter-governmental *Group on Earth Observations* (GEO) to demonstrate how observations from a Global Earth Observation System of Systems (GEOSS) could contribute to improved hydrological ensemble prediction products. HEPEX is one of the GEO Projects (WA-06-02, http://www.earthobservations.org/doc_library/workplan_docs.html).

5. **Next Steps**

Work is now progressing on the first eight test beds, and discussions are underway to begin several others. Supporting data sets and CHPS components are being developed by some of the test bed project scientist. Information about the test beds and algorithm development will be included in annual test bed project reports that are being prepared and will appear on the HEPEX web site.

The next HEPEX workshop is scheduled for June 2007 in Stresa, Italy, where the community will share and debate the latest innovations in ensemble hydrological prediction, and review research progress in the test-beds. We encourage relevant scientist to join us in Italy and at subsequent HEPEX meetings. We particularly would value the contributions from a wider range of hydrological forecast users as well as scientists with expertise in hydrological data assimilation, downscaling, and land-surface modeling.

FOR FURTHER READING

Bandaragoda, C., Tarboton, D. and Woods, R., 2004: Application of topmodel in the Distributed Model Intercomparison Project. *J. Hydrology*, **298**, 178–201.

Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M., & Zhu, Y., 2005: A comparison of the ECMWF, MSC and NCEP Global Ensemble Prediction Systems. *Mon. Wea. Rev*., **133**, 1076,1097.

Clark, M.P., A.G. Slater, A.P. Barrett, B. Rajagopalan, G.J. McCabe, L.E. Hay, and G. H. Leavesley, 2006: Assimilation of snow covered area information into hydrologic and land-surface models. *Advances in Water Resources*, **29**, 1209-1221.

Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts, an important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc*., **87**, 33-46.

Koren, V., Reed, S., Smith, M., Zhang, Z., Seo, D.- J., 2004: Hydrology laboratory research modeling system (HL-RMS) of the US National Weather Service. *J. Hydrology*, **291**, 297–318.

Mitchell, K. E., and coauthors, 2004: The multi-institution North American land data assimilation system (NLDAS): utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res*., **109D**, D07S90. DOI 10.1029/2003JD003823.

National Research Council (Committee on Estimating and Communicating Uncertainty in Weather and Climate Forecasts), 2006: *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts*. National Academy Press, ISBN 0-309-10255-3. 140 pp.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002: Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Wea. Rev*., **130**, 103-114.

Schaake, J., C., and coauthors, 2006: Precipitation and temperature ensemble forecasts from single value forecasts. *Hydrology and Earth System Science, hessd-2006-0094.*

Seo, D. J., V. Koren, N. Cahina, 2003: Real-time variational assimilation of hydrologic and hydrometerological data into operational hydrological forecasting. *J. Hydrometeor*., **4**, 627-641.

Seo, D. J., H. H. Herr, and J. C. Schaake, 2006: A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble Streamflow prediction. *Hydrology and Earth System Science, hessd-2006-0022.*

Seuffert, G., H. Wilker, P. Viterbo, M. Drusch, and J. Mahfouf, 2004: The usage of screen-level parameters and microwave brightness temperature for soil moisture analysis. *J. Hydrometeor*., **5**, 516-531.

Slater, A.G., and M.P. Clark, 2006: Snow data assimilation via an ensemble Kalman filter. *J. Hydrometeor.*, **7**, 478-493.

TIGGE, 2005: THORPEX Interactive Grand Global Ensemble (Summary). Available at http://www.wmo.ch/thorpex/pdf/tigge_summary.pdf

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc*., **83**, 73-83.

List of Figures

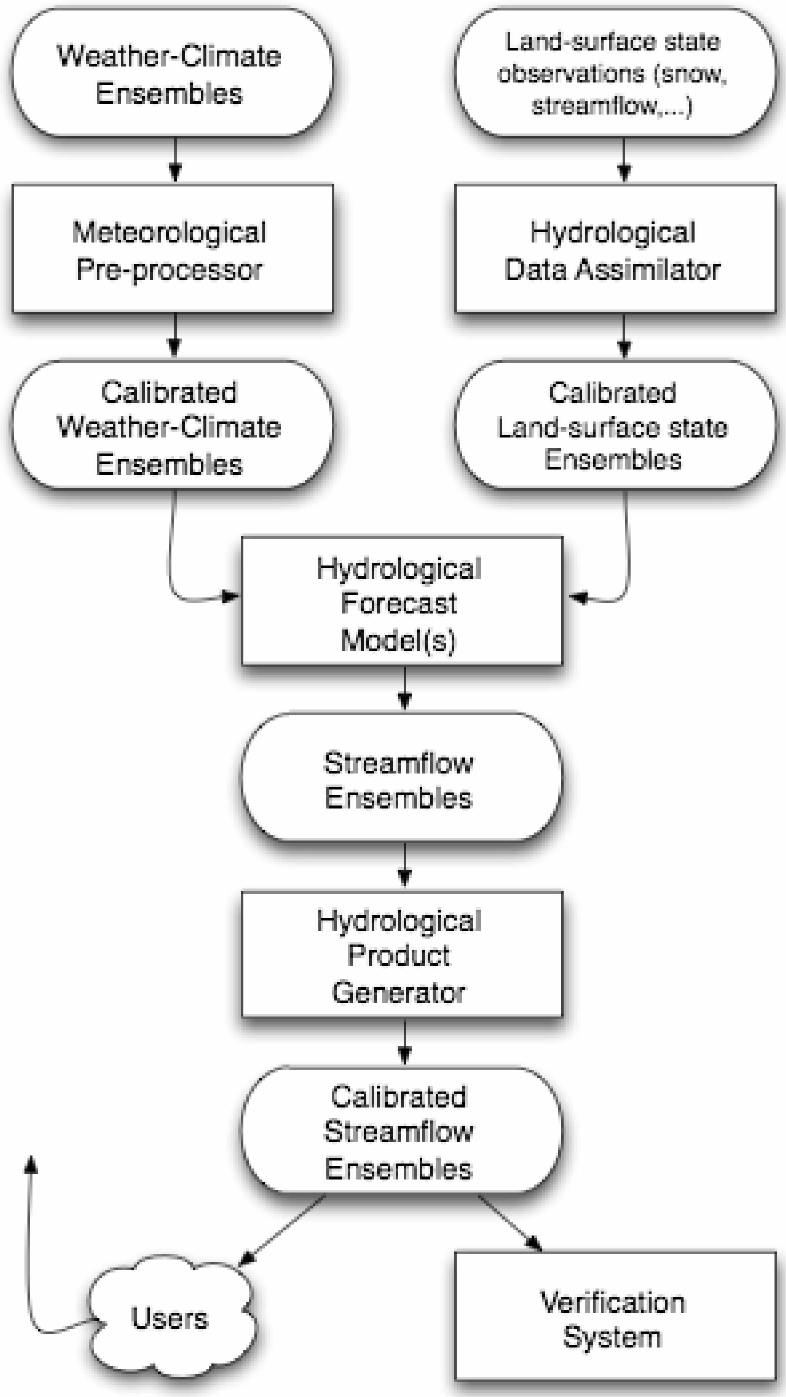

Figure 1: Schematic of the components and information flow in an ensemble hydrological prediction system.

Table 1: HEPEX test bed projects

	Test Bed	Objectives	Contact
1.	Great Lakes	Demonstrate the importance of relatively detailed atmospheric and hydrological modeling for medium-range atmospheric and hydrological forecasting on large basins.	Vincent Fortin vincent.fortin@ec.gc.ca
2.	Bangladesh	Improve operational real-time forecasts of river discharge into Bangladesh at daily, weekly, monthly, and seasonal timescales.	Tom Hopson hopson@ucar.edu
3.	Rio Grande Basin, Brazil	Explore hydrologic predictions driven by forecasts from a global and regional models.	Carlos Tucci tucci@iph.ufrgs.br
4.	Po Basin, Italy	Test simple methods for removing biases from weather forecast models in a region of complex topography, and test flood forecast methods based on precipitation exceedance probabilities.	Jutta Thielen jutta.thielen@jrc.it
5.	Southeast U.S.	Address questions such as: 1. How do we generate skillful and reliable meteorological forcing for seasonal hydrologic forecasting? 2. How do we generate the hydrological ensembles that reflect all sources of uncertainty. 3. How can climate information be be used reliably in seasonal hydrological	Eric Wood efwood@princeton.edu

		forecasting. 4. How do we validate hydrological ensembles for the rare extreme events?	
6.	Western U.S.	Develop hydrological ensemble forecasting techniques that are particular to the orographically complex, snowmelt-driven basins of the western US and British Columbia. Focus on monthly to seasonal lead times.	Frank Weber frank.weber@bchydro.bc.ca Andy Wood aww@hydro.washington.edu
7.	Statistical	1. Identify the space-time scales for for which forecast skill is present for different variables, and develop methods to extract and combine information and different space-time scales. 2. Identify forecast model variables that are the appropriate predictors of sub-gridscale precipitation. 3. Identify the forecast sample size required to calibrate precipitation and temperature forecasts.	Martyn Clark mp.clark@niwa.co.nz John Schaake john.schaake@noaa.gov
8.	Hydrological	Define the advantages and limitations of different methods for characterizing and reducing the uncertainty in hydrological model simulations due to uncertainties in model inputs, parameters, and model structure.	Martyn Clark mp.clark@niwa.co.nz Jasper Vrugt vrugt@lanl.gov